SEQUENCE LISTING

(1) GENERAL INFORMATION:

    (i) APPLICANT:
        (A) NAME: MOGEN INTERNATIONAL NV
        (B) STREET: Einsteinweg 97
        (C) CITY: Leiden
        (E) COUNTRY: The Netherlands
        (F) POSTAL CODE (ZIP): 2233 CB
        (G) TELEPHONE: (31) 71-5258282
        (H) TELEFAX: (31) 71-5221471

    (ii) TITLE OF INVENTION: Enhanced accumulation of trehalose in plants

    (iii) NUMBER OF SEQUENCES: 27

    (iv) COMPUTER READABLE FORM:
        (A) MEDIUM TYPE: Floppy disk
        (B) COMPUTER: IBM PC compatible
        (C) OPERATING SYSTEM: PC-DOS/MS-DOS
        (D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 2621 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA to mRNA

    (iii) HYPOTHETICAL: NO

    (ix) FEATURE:
        (A) NAME/KEY: CDS
        (B) LOCATION: 25..2485
        (D) OTHER INFORMATION: /function= "trehalose phosph.
            synthase and trehalose phosph. phosphatase"
            /product= "bipartite enzyme"

    (ix) FEATURE:
        (A) NAME/KEY: unsure
        (B) LOCATION: 1609..1611

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```
     CTGATCCTGC GGTTTCATCA CAAT ATG ATA CTC TTA CAT CTG ATG CCC CTT          51
                                 Met Ile Leu Leu His Leu Met Pro Leu
  5                               1                   5

     CAG ATG CTC CCA AAT AGG TTG ATT GTC GTA TCG AAT CAG TTA CCC ATA          99
     Gln Met Leu Pro Asn Arg Leu Ile Val Val Ser Asn Gln Leu Pro Ile
      10              15                  20                  25
 10
     ATC GCT AGG CTA AGA CTA ACG ACA ATG GAG GGT CCT TTT GGG ATT TCA         147
     Ile Ala Arg Leu Arg Leu Thr Thr Met Glu Gly Pro Phe Gly Ile Ser
                      30                  35                  40

 15  CTT GGG ACG AGA GTT CGA TTT ACA TGC ACA TCA AAG ATG CAT TAC CCG         195
     Leu Gly Thr Arg Val Arg Phe Thr Cys Thr Ser Lys Met His Tyr Pro
                          45                  50                  55

     CAG CCG TTG AGG TTT TCT ATT CTT GGC GAT CCA CTA AGG GCT GAC GTT         243
 20  Gln Pro Leu Arg Phe Ser Ile Leu Gly Asp Pro Leu Arg Ala Asp Val
                  60                  65                  70

     GGC CCT ACC GAA CAA GAT GAC GTG TCA AAG ACA TTG CTC GAT AGG TTT         291
     Gly Pro Thr Glu Gln Asp Asp Val Ser Lys Thr Leu Leu Asp Arg Phe
 25      75                  80                  85

     AAT TGC GTT GCG GTT TTT GTC CCT ACT TCA AAA TGG GAC CAA TAT TAT         339
     Asn Cys Val Ala Val Phe Val Pro Thr Ser Lys Trp Asp Gln Tyr Tyr
      90                  95                  100                 105
 30
     CAC TGC TTT TGT AAG CAG TAT TTG TGG CCG ATA TTT CAT TAC AAG GTT         387
     His Cys Phe Cys Lys Gln Tyr Leu Trp Pro Ile Phe His Tyr Lys Val
                          110                 115                 120

 35  CCC GCT TCT GAC GTC AAG AGT GTC CCG AAT AGT CGG GAT TCA TGG AAC         435
     Pro Ala Ser Asp Val Lys Ser Val Pro Asn Ser Arg Asp Ser Trp Asn
                      125                 130                 135

     GCT TAT GTT CAC GTG AAC AAA GAG TTT TCC CAG AAG GTG ATG GAG GCA         483
 40  Ala Tyr Val His Val Asn Lys Glu Phe Ser Gln Lys Val Met Glu Ala
              140                 145                 150

     GTA ACC AAT CGT AGC AAT TAT GTA TGG ATA CAT GAC TAC CAT TTA ATG         531
     Val Thr Asn Arg Ser Asn Tyr Val Trp Ile His Asp Tyr His Leu Met
 45      155                 160                 165

     ACG CTA CCG ACT TTC TTG AGG CGG GAT TTT TGT CGT TTT AAA ATC GGT         579
     Thr Leu Pro Thr Phe Leu Arg Arg Asp Phe Cys Arg Phe Lys Ile Gly
     170                 175                 180                 185
 50
     TTT TTT CTG CAT AGC CCG TTT CCT TCC TCG GAG GTT TAC AAG ACC CTA         627
     Phe Phe Leu His Ser Pro Phe Pro Ser Ser Glu Val Tyr Lys Thr Leu
                      190                 195                 200
```

```
      CCA ATG AGA AAC GAG CTC TTG AAG GGT CTG TTA AAT GCT GAT CTT ATC        675
      Pro Met Arg Asn Glu Leu Leu Lys Gly Leu Leu Asn Ala Asp Leu Ile
              205             210             215

  5   GGG TTC CAT ACA TAC GAT TAT GCC CGT CAT TTT CTA ACG TGT TGT AGT        723
      Gly Phe His Thr Tyr Asp Tyr Ala Arg His Phe Leu Thr Cys Cys Ser
              220             225             230

      CGA ATG TTT GGT TTG GAT CAT CAG TTG AAA AGG GGG TAC ATT TTC TTG        771
 10   Arg Met Phe Gly Leu Asp His Gln Leu Lys Arg Gly Tyr Ile Phe Leu
              235             240             245

      GAA TAT AAT GGA AGG AGC ATT GAG ATC AAG ATA AAG GCG AGC GGG ATT        819
      Glu Tyr Asn Gly Arg Ser Ile Glu Ile Lys Ile Lys Ala Ser Gly Ile
 15   250             255             260             265

      CAT GTT GGT CGA ATG GAG TCG TAC TTG AGT CAG CCC GAT ACA AGA TTA        867
      His Val Gly Arg Met Glu Ser Tyr Leu Ser Gln Pro Asp Thr Arg Leu
                      270             275             280
 20
      CAA GTT CAA GAA GTC CAA AAA CGT TCG AAG GAA ATC GTG CTA CTG GGA        915
      Gln Val Gln Glu Val Gln Lys Arg Ser Lys Glu Ile Val Leu Leu Gly
              285             290             295

 25   GTT GAT GAT TTG GAT ATA TTC AAA GGT GTG AAC TTC AAG GTT TTA GCG        963
      Val Asp Asp Leu Asp Ile Phe Lys Gly Val Asn Phe Lys Val Leu Ala
              300             305             310

      TTG GAG AAG TTA CTT AAA TCA CAC CCG AGT TGG CAA GGG CGT GTG GAA       1011
 30   Leu Glu Lys Leu Leu Lys Ser His Pro Ser Trp Gln Gly Arg Val Glu
              315             320             325

      AAG GTG CAA ATC TTG AAT CCT CTG CGC CGT TGC CAA GAC GTC GAT GAG       1059
      Lys Val Gln Ile Leu Asn Pro Leu Arg Arg Cys Gln Asp Val Asp Glu
 35   330             335             340             345

      ATC AAT GCC GAG ATA AGA ACA GTC TGT GAA AGA ATC AAT AAC GAA CTG       1107
      Ile Asn Ala Glu Ile Arg Thr Val Cys Glu Arg Ile Asn Asn Glu Leu
                      350             355             360
 40
      GGA AGC CCG GGA TAC CAG CCC GTT GTG TTA ATT GAT GGG CCC GTT TCG       1155
      Gly Ser Pro Gly Tyr Gln Pro Val Val Leu Ile Asp Gly Pro Val Ser
              365             370             375

 45   TTA AGT GAA AAA GCT GCT TAT TAT GCT ATC GCC GAT ATG GCA ATT GTT       1203
      Leu Ser Glu Lys Ala Ala Tyr Tyr Ala Ile Ala Asp Met Ala Ile Val
              380             385             390

      ACA CCG TTA CGT GAC GGA CTG AAT CTT ATC CCG TAC GAG TAC GTC GTT       1251
 50   Thr Pro Leu Arg Asp Gly Leu Asn Leu Ile Pro Tyr Glu Tyr Val Val
              395             400             405
```

```
    TCC CGA CAA AGT GTT AAT GAC CCA AAT CCC AAT ACT CCA AAA AAG AGC      1299
    Ser Arg Gln Ser Val Asn Asp Pro Asn Pro Asn Thr Pro Lys Lys Ser
    410             415             420             425

 5  ATG CTA GTG GTC TCC GAG TTC ATC GGT GTT TCA CTA TCT TTA ACC GGG      1347
    Met Leu Val Val Ser Glu Phe Ile Gly Val Ser Leu Ser Leu Thr Gly
                    430             435             440

    GCC ATA CGG GTC AAC CCA TGG GAT GAG TTG GAG ACA GCA GAA GCA TTA      1395
10  Ala Ile Arg Val Asn Pro Trp Asp Glu Leu Glu Thr Ala Glu Ala Leu
                445             450             455

    TAC GAC GCA CTC ATG GCT CCT GAT GAC CAT AAA GAA ACC GCC CAC ATG      1443
    Tyr Asp Ala Leu Met Ala Pro Asp Asp His Lys Glu Thr Ala His Met
15          460             465             470

    AAA CAG TAT CAA TAC ATT ATC TCC CAT GAT GTA GCT AAC TGG GCT AGC      1491
    Lys Gln Tyr Gln Tyr Ile Ile Ser His Asp Val Ala Asn Trp Ala Ser
        475             480             485
20
    TTC TTT CAA GAT TTA GAG CAA GCG TGC ATC GAT CAT TCT CGT AAA CGA      1539
    Phe Phe Gln Asp Leu Glu Gln Ala Cys Ile Asp His Ser Arg Lys Arg
    490             495             500             505

25  TGC ATG AAT TTA GGA TTT GGG TTA GAT ACT AGA GTC GTC TTT TTG ATG      1587
    Cys Met Asn Leu Gly Phe Gly Leu Asp Thr Arg Val Val Phe Leu Met
                    510             515             520

    AGA AGT TTA GCA AGT TGG ATA AAG ATG TCT TGG AAG AAT GCT TAT TCC      1635
30  Arg Ser Leu Ala Ser Trp Ile Lys Met Ser Trp Lys Asn Ala Tyr Ser
                525             530             535

    ATG GCT CAA AAT CGG GCC ATA CTT TTG GAC TAT GAC GGC ACT GTT ACT      1683
    Met Ala Gln Asn Arg Ala Ile Leu Leu Asp Tyr Asp Gly Thr Val Thr
35          540             545             550

    CCA TCT ATC AGT AAA TCT CCA ACT GAA GCT GTT ATC TCC ATG ATC AAC      1731
    Pro Ser Ile Ser Lys Ser Pro Thr Glu Ala Val Ile Ser Met Ile Asn
        555             560             565
40
    AAA CTG TGC AAT GAT CCA AAG AAC ATG GTG TTC ATC GTT AGT GGA CGC      1779
    Lys Leu Cys Asn Asp Pro Lys Asn Met Val Phe Ile Val Ser Gly Arg
    570             575             580             585

45  AGT AGA GAG AAA ATC TTG GCA GTT GGT TCG GCG CGT GTG AGA ACC CGC      1827
    Ser Arg Glu Lys Ile Leu Ala Val Gly Ser Ala Arg Val Arg Thr Arg
                590             595             600

    CAT TGC ACT GAG CAC GGA TAC TTT ATA AGG TGG GCG GGT GAT CAA GAA      1875
50  His Cys Thr Glu His Gly Tyr Phe Ile Arg Trp Ala Gly Asp Gln Glu
                605             610             615
```

```
      TGG GAA ACG TGC GCA CGT GAG AAT AAT GTC GGG TGG ATG GAT GGA AAT          1923
      Trp Glu Thr Cys Ala Arg Glu Asn Asn Val Gly Trp Met Asp Gly Asn
          620                 625                 630

   5  CTG AGG CCG GTT ATG AAT CTT TAT ACA GAA ACT ACT GAC GGT TCG TAT          1971
      Leu Arg Pro Val Met Asn Leu Tyr Thr Glu Thr Thr Asp Gly Ser Tyr
          635                 640                 645

      ATT GAA AAG AAA GAA ACT GCA ATG GTT TGG CAC TAT GAA GAT GCT GAT          2019
  10  Ile Glu Lys Lys Glu Thr Ala Met Val Trp His Tyr Glu Asp Ala Asp
          650                 655                 660                 665

      AAA GAT CTT GGG TTG GAG CAG GCT AAG GAA CTG TTG GAC CAT CTT GAA          2067
      Lys Asp Leu Gly Leu Glu Gln Ala Lys Glu Leu Leu Asp His Leu Glu
  15                  670                 675                 680

      AAC GTG CTC GCT AAT GAG CCC GTT GGA GTG AAT CGA ACA GGT CAA TAC          2115
      Asn Val Leu Ala Asn Glu Pro Val Gly Val Asn Arg Thr Gly Gln Tyr
                      685                 690                 695
  20
      ATT GTA GAA GTT AAA CCA CAG TCC CCC ATT AAT TAC CTT CTT GTT ATG          2163
      Ile Val Glu Val Lys Pro Gln Ser Pro Ile Asn Tyr Leu Leu Val Met
          700                 705                 710

  25  ACA TTC ATA GGC ACT GAT TGT AGA ATC TTT AAC TTA AAT TTC TTT AAA          2211
      Thr Phe Ile Gly Thr Asp Cys Arg Ile Phe Asn Leu Asn Phe Phe Lys
          715                 720                 725

      TAT GAA TGC AAT TAT AGG GGG TCA CTA AAA GGT ATA GTT GCA GAG AAG          2259
  30  Tyr Glu Cys Asn Tyr Arg Gly Ser Leu Lys Gly Ile Val Ala Glu Lys
          730                 735                 740                 745

      ATT TTT GCG TTC ATG GCT AAA AAG GGA AAA CAG GCT GAT TTC GTG TTG          2307
      Ile Phe Ala Phe Met Ala Lys Lys Gly Lys Gln Ala Asp Phe Val Leu
  35                  750                 755                 760

      ACG TTG AAT GAT AGA AGT GAT GAA GAC ATG TTT GTG GCC ATT GGG GAT          2355
      Thr Leu Asn Asp Arg Ser Asp Glu Asp Met Phe Val Ala Ile Gly Asp
                      765                 770                 775
  40
      GGA ATA AAA AAG GGT CGG ATA ACT AAC AAC AAT TCA GTG TTT ACA TGC          2403
      Gly Ile Lys Lys Gly Arg Ile Thr Asn Asn Asn Ser Val Phe Thr Cys
                      780                 785                 790

  45  GTA GTG GGA GAG AAA CCG AGT GCA GCT GAG TAC TTT TTA AAT GAT GTC          2451
      Val Val Gly Glu Lys Pro Ser Ala Ala Glu Tyr Phe Leu Asn Asp Val
          795                 800                 805

      TCG AGA AGC TCC GGG TGT CTC AGC AAC CAA GGA T GATCCGGAAG              2495
  50  Ser Arg Ser Ser Gly Cys Leu Ser Asn Gln Gly
          810                 815                 820

      CTTCTCGTGA TCTTTATGAG TTAAAAGTTT TCGACTTTTT CTTCATCAAG ATTCATGGGA        2555
```

AAGTTGTTCA ATATGAACTT GTGTTCTTGG TTCTGGATTT TAGGGAGTCT ATGGATATAA   2615

CATTTC   2621

(2) INFORMATION FOR SEQ ID NO: 2:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 820 amino acids
       (B) TYPE: amino acid
       (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```
Met Ile Leu Leu His Leu Met Pro Leu Gln Met Leu Pro Asn Arg Leu
 1               5                  10                  15

Ile Val Val Ser Asn Gln Leu Pro Ile Ile Ala Arg Leu Arg Leu Thr
            20                  25                  30

Thr Met Glu Gly Pro Phe Gly Ile Ser Leu Gly Thr Arg Val Arg Phe
        35                  40                  45

Thr Cys Thr Ser Lys Met His Tyr Pro Gln Pro Leu Arg Phe Ser Ile
        50                  55                  60

Leu Gly Asp Pro Leu Arg Ala Asp Val Gly Pro Thr Glu Gln Asp Asp
    65                  70                  75                  80

Val Ser Lys Thr Leu Leu Asp Arg Phe Asn Cys Val Ala Val Phe Val
                85                  90                  95

Pro Thr Ser Lys Trp Asp Gln Tyr Tyr His Cys Phe Cys Lys Gln Tyr
            100                 105                 110

Leu Trp Pro Ile Phe His Tyr Lys Val Pro Ala Ser Asp Val Lys Ser
            115                 120                 125

Val Pro Asn Ser Arg Asp Ser Trp Asn Ala Tyr Val His Val Asn Lys
            130                 135                 140

Glu Phe Ser Gln Lys Val Met Glu Ala Val Thr Asn Arg Ser Asn Tyr
145                 150                 155                 160

Val Trp Ile His Asp Tyr His Leu Met Thr Leu Pro Thr Phe Leu Arg
                165                 170                 175

Arg Asp Phe Cys Arg Phe Lys Ile Gly Phe Phe Leu His Ser Pro Phe
            180                 185                 190

Pro Ser Ser Glu Val Tyr Lys Thr Leu Pro Met Arg Asn Glu Leu Leu
            195                 200                 205
```

```
      Lys Gly Leu Leu Asn Ala Asp Leu Ile Gly Phe His Thr Tyr Asp Tyr
          210             215             220
  5   Ala Arg His Phe Leu Thr Cys Cys Ser Arg Met Phe Gly Leu Asp His
          225             230             235             240
      Gln Leu Lys Arg Gly Tyr Ile Phe Leu Glu Tyr Asn Gly Arg Ser Ile
                  245             250             255
 10
      Glu Ile Lys Ile Lys Ala Ser Gly Ile His Val Gly Arg Met Glu Ser
                  260             265             270
      Tyr Leu Ser Gln Pro Asp Thr Arg Leu Gln Val Gln Glu Val Gln Lys
 15           275             280             285
      Arg Ser Lys Glu Ile Val Leu Leu Gly Val Asp Asp Leu Asp Ile Phe
              290             295             300
 20   Lys Gly Val Asn Phe Lys Val Leu Ala Leu Glu Lys Leu Leu Lys Ser
          305             310             315             320
      His Pro Ser Trp Gln Gly Arg Val Glu Lys Val Gln Ile Leu Asn Pro
                  325             330             335
 25
     Leu Arg Arg Cys Gln Asp Val Asp Glu Ile Asn Ala Glu Ile Arg Thr
              340             345             350
      Val Cys Glu Arg Ile Asn Asn Glu Leu Gly Ser Pro Gly Tyr Gln Pro
 30           355             360             365
      Val Val Leu Ile Asp Gly Pro Val Ser Leu Ser Glu Lys Ala Ala Tyr
          370             375             380
 35   Tyr Ala Ile Ala Asp Met Ala Ile Val Thr Pro Leu Arg Asp Gly Leu
          385             390             395             400
      Asn Leu Ile Pro Tyr Glu Tyr Val Val Ser Arg Gln Ser Val Asn Asp
                  405             410             415
 40
      Pro Asn Pro Asn Thr Pro Lys Lys Ser Met Leu Val Val Ser Glu Phe
                  420             425             430
      Ile Gly Val Ser Leu Ser Leu Thr Gly Ala Ile Arg Val Asn Pro Trp
 45           435             440             445
      Asp Glu Leu Glu Thr Ala Glu Ala Leu Tyr Asp Ala Leu Met Ala Pro
          450             455             460
 50   Asp Asp His Lys Glu Thr Ala His Met Lys Gln Tyr Gln Tyr Ile Ile
          465             470             475             480
      Ser His Asp Val Ala Asn Trp Ala Ser Phe Phe Gln Asp Leu Glu Gln
                  485             490             495
```

```
   Ala Cys Ile Asp His Ser Arg Lys Arg Cys Met Asn Leu Gly Phe Gly
           500             505             510

 5 Leu Asp Thr Arg Val Val Phe Leu Met Arg Ser Leu Ala S r Trp Ile
           515             520             525

   Lys Met Ser Trp Lys Asn Ala Tyr Ser Met Ala Gln Asn Arg Ala Ile
           530             535             540
10
   Leu Leu Asp Tyr Asp Gly Thr Val Thr Pro Ser Ile Ser Lys Ser Pro
   545             550             555             560

   Thr Glu Ala Val Ile Ser Met Ile Asn Lys Leu Cys Asn Asp Pro Lys
15             565             570             575

   Asn Met Val Phe Ile Val Ser Gly Arg Ser Arg Glu Lys Ile Leu Ala
                   580             585             590

20 Val Gly Ser Ala Arg Val Arg Thr Arg His Cys Thr Glu His Gly Tyr
           595             600             605

   Phe Ile Arg Trp Ala Gly Asp Gln Glu Trp Glu Thr Cys Ala Arg Glu
           610             615             620
25
   Asn Asn Val Gly Trp Met Asp Gly Asn Leu Arg Pro Val Met Asn Leu
   625             630             635             640

   Tyr Thr Glu Thr Thr Asp Gly Ser Tyr Ile Glu Lys Lys Glu Thr Ala
                   645             650             655
30
   Met Val Trp His Tyr Glu Asp Ala Asp Lys Asp Leu Gly Leu Glu Gln
           660             665             670

35 Ala Lys Glu Leu Leu Asp His Leu Glu Asn Val Leu Ala Asn Glu Pro
           675             680             685

   Val Gly Val Asn Arg Thr Gly Gln Tyr Ile Val Glu Val Lys Pro Gln
           690             695             700
40
   Ser Pro Ile Asn Tyr Leu Leu Val Met Thr Phe Ile Gly Thr Asp Cys
   705             710             715             720

   Arg Ile Phe Asn Leu Asn Phe Phe Lys Tyr Glu Cys Asn Tyr Arg Gly
45                 725             730             735

   Ser Leu Lys Gly Ile Val Ala Glu Lys Ile Phe Ala Phe Met Ala Lys
                   740             745             750

50 Lys Gly Lys Gln Ala Asp Phe Val Leu Thr Leu Asn Asp Arg Ser Asp
           755             760             765

   Glu Asp Met Phe Val Ala Ile Gly Asp Gly Ile Lys Lys Gly Arg Ile
           770             775             780
```

```
    Thr Asn Asn Asn Ser Val Phe Thr Cys Val Val Gly Glu Lys Pro Ser
    785             790             795             800
```

5

```
    Ala Ala Glu Tyr Phe Leu Asn Asp Val Ser Arg Ser Ser Gly Cys Leu
                805             810             815
```

```
    Ser Asn Gln Gly
                820
```

10

(2) INFORMATION FOR SEQ ID NO: 3:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 25 base pairs

15
       (B) TYPE: nucleic acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

20

    (iii) HYPOTHETICAL: YES

    (iii) ANTI-SENSE: NO

25

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

AAGCTTATGT TGCCATATAG AGTAG   25

30

(2) INFORMATION FOR SEQ ID NO: 4:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 24 base pairs

35
       (B) TYPE: nucleic acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

40

    (iii) HYPOTHETICAL: YES

    (iii) ANTI-SENSE: NO

45

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

GTAGTTGCCA TGGTGCAAAT GTTC   24

50

(2) INFORMATION FOR SEQ ID NO: 5:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 20 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (iii) HYPOTHETICAL: YES

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

AGCTCTGCAG TGAGGTACCA   20

(2) INFORMATION FOR SEQ ID NO: 6:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 20 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (iii) HYPOTHETICAL: YES

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

GACGTCACTC CATGGTTCGA   20

(2) INFORMATION FOR SEQ ID NO: 7:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 24 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (iii) HYPOTHETICAL: YES

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

GTACCCTGCA GTGTGACCCT AGAC   24

(2) INFORMATION FOR SEQ ID NO: 8:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 21 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (iii) HYPOTHETICAL: YES

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

TCGATTCATA GAAGCTTAGA T   21

(2) INFORMATION FOR SEQ ID NO: 9:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 2207 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: double
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA to mRNA

    (iii) HYPOTHETICAL: NO

    (iii) ANTI-SENSE: NO

    (vi) ORIGINAL SOURCE:
        (A) ORGANISM: Solanum tuberosum
        (B) STRAIN: Kardal

    (ix) FEATURE:
        (A) NAME/KEY: CDS
        (B) LOCATION: 161..1906

    (ix) FEATURE:
        (A) NAME/KEY: misc_feature
        (B) LOCATION: 842..850
        (D) OTHER INFORMATION: /function= "putative
            glycosylationsite"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

CTTTTCTGAG TAATAACATA GGCATTGATT TTTTTTCAAT TAATAACACC TGCAAACATT   60

CCCATTGCCG GCATTCTCTG TTCTTACAAA AAAAAACATT TTTTTGTTCA CATAAATTAG   120

```
TTATGGCATC AGTATTGAAC CCTTTAACTT GTTATACAAT ATG GGT AAA GCT ATA          175
                                             Met Gly Lys Ala Ile
                                             1                 5

  5 ATT TTT ATG ATT TTT ACT ATG TCT ATG AAT ATG ATT AAA GCT GAA ACT          223
    Ile Phe Met Ile Phe Thr Met Ser Met Asn Met Ile Lys Ala Glu Thr
                 10              15              20

    TGC AAA TCC ATT GAT AAG GGT CCT GTA ATC CCA ACA ACC CCT TTA GTG          271
 10 Cys Lys Ser Ile Asp Lys Gly Pro Val Ile Pro Thr Thr Pro Leu Val
                 25              30              35

    ATT TTT CTT GAA AAA GTT CAA GAA GCT GCT CTT CAA ACT TAT GGC CAT          319
    Ile Phe Leu Glu Lys Val Gln Glu Ala Ala Leu Gln Thr Tyr Gly His
 15         40              45              50

    AAA GGG TTT GAT GCT AAA CTG TTT GTT GAT ATG TCA CTG AGA GAG AGT          367
    Lys Gly Phe Asp Ala Lys Leu Phe Val Asp Met Ser Leu Arg Glu Ser
         55              60              65
 20

    CTT TCA GAA ACA GTT GAA GCT TTT AAT AAG CTT CCA AGA GTT GTG AAT          415
    Leu Ser Glu Thr Val Glu Ala Phe Asn Lys Leu Pro Arg Val Val Asn
            70              75              80              85

 25 GGT TCA ATA TCA AAA AGT GAT TTG GAT GGT TTT ATA GGT AGT TAC TTG          463
    Gly Ser Ile Ser Lys Ser Asp Leu Asp Gly Phe Ile Gly Ser Tyr Leu
                 90              95              100

    AGT AGT CCT GAT AAG GAT TTG GTT TAT GTT GAG CCT ATG GAT TTT GTG          511
 30 Ser Ser Pro Asp Lys Asp Leu Val Tyr Val Glu Pro Met Asp Phe Val
                 105             110             115

    GCT GAG CCT GAA GGC TTT TTG CCA AAG GTG AAG AAT TCT GAG GTG AGG          559
    Ala Glu Pro Glu Gly Phe Leu Pro Lys Val Lys Asn Ser Glu Val Arg
 35         120             125             130

    GCA TGG GCA TTG GAG GTG CAT TCA CTT TGG AAG AAT TTA AGT AGG AAA          607
    Ala Trp Ala Leu Glu Val His Ser Leu Trp Lys Asn Leu Ser Arg Lys
         135             140             145
 40

    GTG GCT GAT CAT GTA TTG GAA AAA CCA GAG TTG TAT ACT TTG CTT CCA          655
    Val Ala Asp His Val Leu Glu Lys Pro Glu Leu Tyr Thr Leu Leu Pro
    150             155             160             165

 45 TTG AAA AAT CCA GTT ATT ATA CCG GGA TCG CGT TTT AAG GAG GTT TAT          703
    Leu Lys Asn Pro Val Ile Ile Pro Gly Ser Arg Phe Lys Glu Val Tyr
                 170             175             180

    TAT TGG GAT TCT TAT TGG GTA ATA AGG GGT TTG TTA GCA AGC AAA ATG          751
 50 Tyr Trp Asp Ser Tyr Trp Val Ile Arg Gly Leu Leu Ala Ser Lys Met
                 185             190             195
```

```
    TAT GAA ACT GCA AAA GGG ATT GTG ACT AAT CTG GTT TCT CTG ATA GAT    799
    Tyr Glu Thr Ala Lys Gly Ile Val Thr Asn Leu Val Ser Leu Ile Asp
            200                 205                 210

  5 CAA TTT GGT TAT GTT CTT AAC GGT GCA AGA GCA TAC TAC AGT AAC AGA    847
    Gln Phe Gly Tyr Val Leu Asn Gly Ala Arg Ala Tyr Tyr Ser Asn Arg
            215                 220                 225

    AGT CAG CCT CCT GTC CTG GCC ACG ATG ATT GTT GAC ATA TTC AAT CAG    895
 10 Ser Gln Pro Pro Val Leu Ala Thr Met Ile Val Asp Ile Phe Asn Gln
    230                 235                 240                 245

    ACA GGT GAT TTA AAT TTG GTT AGA AGA TCC CTT CCT GCT TTG CTC AAG    943
    Thr Gly Asp Leu Asn Leu Val Arg Arg Ser Leu Pro Ala Leu Leu Lys
 15                 250                 255                 260

    GAG AAT CAT TTT TGG AAT TCA GGA ATA CAT AAG GTG ACT ATT CAA GAT    991
    Glu Asn His Phe Trp Asn Ser Gly Ile His Lys Val Thr Ile Gln Asp
                265                 270                 275
 20
    GCT CAG GGA TCA AAC CAC AGC TTG AGT CGG TAC TAT GCT ATG TGG AAT   1039
    Ala Gln Gly Ser Asn His Ser Leu Ser Arg Tyr Tyr Ala Met Trp Asn
                280                 285                 290

 25 AAG CCC CGT CCA GAA TCG TCA ACT ATA GAC AGT GAA ACA GCT TCC GTA   1087
    Lys Pro Arg Pro Glu Ser Ser Thr Ile Asp Ser Glu Thr Ala Ser Val
            295                 300                 305

    CTC CCA AAT ATA TGT GAA AAA AGA GAA TTA TAC CGT GAA CTG GCA TCA   1135
 30 Leu Pro Asn Ile Cys Glu Lys Arg Glu Leu Tyr Arg Glu Leu Ala Ser
    310                 315                 320                 325

    GCT GCT GAA AGT GGA TGG GAT TTC AGT TCA AGA TGG ATG AGC AAC GGA   1183
    Ala Ala Glu Ser Gly Trp Asp Phe Ser Ser Arg Trp Met Ser Asn Gly
                330                 335                 340
 35
    TCT GAT CTG ACA ACA ACT AGT ACA ACA TCA ATT CTA CCA GTT GAT TTG   1231
    Ser Asp Leu Thr Thr Thr Ser Thr Thr Ser Ile Leu Pro Val Asp Leu
                345                 350                 355
 40
    AAT GCA TTC CTT CTG AAG ATG GAA CTT GAC ATT GCC TTT CTA GCA AAT   1279
    Asn Ala Phe Leu Leu Lys Met Glu Leu Asp Ile Ala Phe Leu Ala Asn
            360                 365                 370

 45 CTT GTT GGA GAA AGT AGC ACG GCT TCA CAT TTT ACA GAA GCT GCT CAA   1327
    Leu Val Gly Glu Ser Ser Thr Ala Ser His Phe Thr Glu Ala Ala Gln
            375                 380                 385

    AAT AGA CAG AAG GCT ATA AAC TGT ATC TTT TGG AAC GCA GAG ATG GGG   1375
 50 Asn Arg Gln Lys Ala Ile Asn Cys Ile Phe Trp Asn Ala Glu Met Gly
            390                 395                 400                 405
```

```
   CAA TGG CTT GAT TAC TGG CTT ACC AAC AGC GAC ACA TCT GAG GAT ATT      1423
   Gln Trp Leu Asp Tyr Trp Leu Thr Asn Ser Asp Thr Ser Glu Asp Ile
                   410             415             420

 5 TAT AAA TGG GAA GAT TTG CAC CAG AAC AAG AAG TCA TTT GCC TCT AAT      1471
   Tyr Lys Trp Glu Asp Leu His Gln Asn Lys Lys Ser Phe Ala Ser Asn
                   425             430             435

   TTT GTT CCG CTG TGG ACT GAA ATT TCT TGT TCA GAT AAT AAT ATC ACA      1519
10 Phe Val Pro Leu Trp Thr Glu Ile Ser Cys Ser Asp Asn Asn Ile Thr
                   440             445             450

   ACT CAG AAA GTA GTT CAA AGT CTC ATG AGC TCG GGC TTG CTT CAG CCT      1567
   Thr Gln Lys Val Val Gln Ser Leu Met Ser Ser Gly Leu Leu Gln Pro
15         455             460             465

   GCA GGG ATT GCA ATG ACC TTG TCT AAT ACT GGA CAG CAA TGG GAT TTT      1615
   Ala Gly Ile Ala Met Thr Leu Ser Asn Thr Gly Gln Gln Trp Asp Phe
   470             475             480             485
20
   CCG AAT GGT TGG CCC CCC CTT CAA CAC ATA ATC ATT GAA GGT CTC TTA      1663
   Pro Asn Gly Trp Pro Pro Leu Gln His Ile Ile Ile Glu Gly Leu Leu
                   490             495             500

25 AGG TCT GGA CTA GAA GAG GCA AGA ACC TTA GCA AAA GAC ATT GCT ATT      1711
   Arg Ser Gly Leu Glu Glu Ala Arg Thr Leu Ala Lys Asp Ile Ala Ile
                   505             510             515

   CGC TGG TTA AGA ACT AAC TAT GTG ACT TAC AAG AAA ACC GGT GCT ATG      1759
30 Arg Trp Leu Arg Thr Asn Tyr Val Thr Tyr Lys Lys Thr Gly Ala Met
                   520             525             530

   TAT GAA AAA TAT GAT GTC ACA AAA TGT GGA GCA TAT GGA GGT GGT GGT      1807
   Tyr Glu Lys Tyr Asp Val Thr Lys Cys Gly Ala Tyr Gly Gly Gly Gly
35         535             540             545

   GAA TAT ATG TCC CAA ACG GGT TTC GGA TGG TCA AAT GGC GTT GTA CTG      1855
   Glu Tyr Met Ser Gln Thr Gly Phe Gly Trp Ser Asn Gly Val Val Leu
   550             555             560             565
40
   GCA CTT CTA GAG GAA TTT GGA TGG CCT GAA GAT TTG AAG ATT GAT TGC      1903
   Ala Leu Leu Glu Glu Phe Gly Trp Pro Glu Asp Leu Lys Ile Asp Cys
                   570             575             580

45 TAATGAGCAA GTAGAAAAGC CAAATGAAAC ATCATTGAGT TTTATTTTCT TCTTTTGTTA   1963

   AAATAAGCTG CAATGGTTTG CTGATAGTTT ATGTTTTGTA TTACTATTTC ATAAGGTTTT   2023

   TGTACCATAT CAAGTGATAT TACCATGAAC TATGTCGTTC GGACTCTTCA AATCGGATTT   2083
50
   TGCAAAAATA ATGCAGTTTT GGAGAATCCG ATAACATAGA CCATGTATGG ATCTAAATTG   2143

   TAAACAGCTT ACTATATTAA GTAAAGAAA GATGATTCCT CTGCTTTAAA AAAAAAAAAA    2203
```

(2) INFORMATION FOR SEQ ID NO: 10:

     (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 581 amino acids
        (B) TYPE: amino acid
        (D) TOPOLOGY: linear

     (ii) MOLECULE TYPE: protein

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

```
Met Gly Lys Ala Ile Ile Phe Met Ile Phe Thr Met Ser Met Asn Met
 1               5                  10                  15

Ile Lys Ala Glu Thr Cys Lys Ser Ile Asp Lys Gly Pro Val Ile Pro
                20                  25                  30

Thr Thr Pro Leu Val Ile Phe Leu Glu Lys Val Gln Glu Ala Ala Leu
                35                  40                  45

Gln Thr Tyr Gly His Lys Gly Phe Asp Ala Lys Leu Phe Val Asp Met
            50                  55                  60

Ser Leu Arg Glu Ser Leu Ser Glu Thr Val Glu Ala Phe Asn Lys Leu
 65                  70                  75                  80

Pro Arg Val Val Asn Gly Ser Ile Ser Lys Ser Asp Leu Asp Gly Phe
                85                  90                  95

Ile Gly Ser Tyr Leu Ser Ser Pro Asp Lys Asp Leu Val Tyr Val Glu
                100                 105                 110

Pro Met Asp Phe Val Ala Glu Pro Glu Gly Phe Leu Pro Lys Val Lys
            115                 120                 125

Asn Ser Glu Val Arg Ala Trp Ala Leu Glu Val His Ser Leu Trp Lys
        130                 135                 140

Asn Leu Ser Arg Lys Val Ala Asp His Val Leu Glu Lys Pro Glu Leu
145                 150                 155                 160

Tyr Thr Leu Leu Pro Leu Lys Asn Pro Val Ile Ile Pro Gly Ser Arg
                165                 170                 175

Phe Lys Glu Val Tyr Tyr Trp Asp Ser Tyr Trp Val Ile Arg Gly Leu
                180                 185                 190

Leu Ala Ser Lys Met Tyr Glu Thr Ala Lys Gly Ile Val Thr Asn Leu
            195                 200                 205
```

```
    Val Ser Leu Ile Asp Gln Phe Gly Tyr Val Leu Asn Gly Ala Arg Ala
        210                 215             220

    Tyr Tyr Ser Asn Arg Ser Gln Pro Pro Val Leu Ala Thr Met Ile Val
  5 225                 230             235                 240

    Asp Ile Phe Asn Gln Thr Gly Asp Leu Asn Leu Val Arg Arg Ser Leu
                    245             250                 255

 10 Pro Ala Leu Leu Lys Glu Asn His Phe Trp Asn Ser Gly Ile His Lys
                    260             265             270

    Val Thr Ile Gln Asp Ala Gln Gly Ser Asn His Ser Leu Ser Arg Tyr
                275             280             285
 15
    Tyr Ala Met Trp Asn Lys Pro Arg Pro Glu Ser Ser Thr Ile Asp Ser
        290             295             300

    Glu Thr Ala Ser Val Leu Pro Asn Ile Cys Glu Lys Arg Glu Leu Tyr
 20 305             310             315             320

    Arg Glu Leu Ala Ser Ala Ala Glu Ser Gly Trp Asp Phe Ser Ser Arg
                325             330             335

 25 Trp Met Ser Asn Gly Ser Asp Leu Thr Thr Thr Ser Thr Thr Ser Ile
                340             345             350

    Leu Pro Val Asp Leu Asn Ala Phe Leu Leu Lys Met Glu Leu Asp Ile
        355             360             365
 30
    Ala Phe Leu Ala Asn Leu Val Gly Glu Ser Ser Thr Ala Ser His Phe
        370             375             380

    Thr Glu Ala Ala Gln Asn Arg Gln Lys Ala Ile Asn Cys Ile Phe Trp
 35 385             390             395             400

    Asn Ala Glu Met Gly Gln Trp Leu Asp Tyr Trp Leu Thr Asn Ser Asp
                405             410             415

 40 Thr Ser Glu Asp Ile Tyr Lys Trp Glu Asp Leu His Gln Asn Lys Lys
                420             425             430

    Ser Phe Ala Ser Asn Phe Val Pro Leu Trp Thr Glu Ile Ser Cys Ser
        435             440             445
 45
    Asp Asn Asn Ile Thr Thr Gln Lys Val Val Gln Ser Leu Met Ser Ser
        450             455             460

    Gly Leu Leu Gln Pro Ala Gly Ile Ala Met Thr Leu Ser Asn Thr Gly
 50 465             470             475             480

    Gln Gln Trp Asp Phe Pro Asn Gly Trp Pro Pro Leu Gln His Ile Ile
                485             490             495
```

Il Glu Gly Leu Leu Arg Ser Gly Leu Glu Glu Ala Arg Thr Leu Ala
              500   505   510

Lys Asp Ile Ala Ile Arg Trp Leu Arg Thr Asn Tyr Val Thr Tyr Lys
              515   520   525

Lys Thr Gly Ala Met Tyr Glu Lys Tyr Asp Val Thr Lys Cys Gly Ala
    530   535   540

Tyr Gly Gly Gly Gly Glu Tyr Met Ser Gln Thr Gly Phe Gly Trp Ser
545   550   555   560

Asn Gly Val Val Leu Ala Leu Leu Glu Glu Phe Gly Trp Pro Glu Asp
              565   570   575

Leu Lys Ile Asp Cys
              580

(2) INFORMATION FOR SEQ ID NO: 11:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 33 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (iii) HYPOTHETICAL: YES

    (ix) FEATURE:
        (A) NAME/KEY: modified_base
        (B) LOCATION: 6
        (D) OTHER INFORMATION: /mod_base= i

    (ix) FEATURE:
        (A) NAME/KEY: modified_base
        (B) LOCATION: 15
        (D) OTHER INFORMATION: /mod_base= i

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

GGYGGNMGMT TYRWNGARKT MTAYKRYTGG GAC   33

(2) INFORMATION FOR SEQ ID NO: 12:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 26 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: DNA (genomic)

        (iii) HYPOTHETICAL: YES

        (ix) FEATURE:
             (A) NAME/KEY: modified_base
             (B) LOCATION: 3
             (D) OTHER INFORMATION: /mod_base= i

        (ix) FEATURE:
             (A) NAME/KEY: modified_base
             (B) LOCATION: 6
             (D) OTHER INFORMATION: /mod_base= i

        (ix) FEATURE:
             (A) NAME/KEY: modified_base
             (B) LOCATION: 9
             (D) OTHER INFORMATION: /mod_base= i

        (ix) FEATURE:
             (A) NAME/KEY: modified_base
             (B) LOCATION: 12
             (D) OTHER INFORMATION: /mod_base= i

        (ix) FEATURE:
             (A) NAME/KEY: modified_base
             (B) LOCATION: 15
             (D) OTHER INFORMATION: /mod_base= i

        (ix) FEATURE:
             (A) NAME/KEY: modified_base
             (B) LOCATION: 21
             (D) OTHER INFORMATION: /mod_base= i

        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

GTNCCNGGNG GNCGNTTYRW NGARKT   26

    (2) INFORMATION FOR SEQ ID NO: 13:

        (i) SEQUENCE CHARACTERISTICS:
             (A) LENGTH: 26 base pairs
             (B) TYPE: nucleic acid
             (C) STRANDEDNESS: single
             (D) TOPOLOGY: linear

        (ii) MOLECULE TYPE: DNA (genomic)

        (iii) HYPOTHETICAL: YES

```
     (ix) FEATURE:
          (A) NAME/KEY: modified_base
          (B) LOCATION: 3
          (D) OTHER INFORMATION: /mod_base= i

     (ix) FEATURE:
          (A) NAME/KEY: modified_base
          (B) LOCATION: 9
          (D) OTHER INFORMATION: /mod_base= i

     (ix) FEATURE:
          (A) NAME/KEY: modified_base
          (B) LOCATION: 12
          (D) OTHER INFORMATION: /mod_base= i

     (ix) FEATURE:
          (A) NAME/KEY: modified_base
          (B) LOCATION: 15
          (D) OTHER INFORMATION: /mod_base= i

     (ix) FEATURE:
          (A) NAME/KEY: modified_base
          (B) LOCATION: 18
          (D) OTHER INFORMATION: /mod_base= i


     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

GGNGGYTGNS WNCGNYRNAG RTARTA                                    26

(2) INFORMATION FOR SEQ ID NO: 14:

     (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 24 base pairs
          (B) TYPE: nucleic acid
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

   (iii) HYPOTHETICAL: YES


     (ix) FEATURE:
          (A) NAME/KEY: modified_base
          (B) LOCATION: 1
          (D) OTHER INFORMATION: /mod_base= i

     (ix) FEATURE:
          (A) NAME/KEY: modified_base
          (B) LOCATION: 7
          (D) OTHER INFORMATION: /mod_base= i
```

     (ix) FEATURE:
          (A) NAME/KEY: modified_base
          (B) LOCATION: 19
          (D) OTHER INFORMATION: /mod_base= i

     (ix) FEATURE:
          (A) NAME/KEY: modified_base
          (B) LOCATION: 22
          (D) OTHER INFORMATION: /mod_base= i

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

NSCRTTNRYC CATCCRAANC CNTC

(2) INFORMATION FOR SEQ ID NO: 15:

     (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 20 base pairs
          (B) TYPE: nucleic acid
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (iii) HYPOTHETICAL: NO

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

CGAAACGGGC CCATCAATTA

(2) INFORMATION FOR SEQ ID NO: 16:

     (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 20 base pairs
          (B) TYPE: nucleic acid
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (iii) HYPOTHETICAL: NO

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

TCGATGAGAT CAATGCCGAG

24

20

20

(2) INFORMATION FOR SEQ ID NO: 17:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 27 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: cDNA

  (iii) HYPOTHETICAL: NO

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

CCATCCTAAT ACGACTCACT ATAGGGC   27

(2) INFORMATION FOR SEQ ID NO: 18:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 20 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: cDNA

  (iii) HYPOTHETICAL: NO

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

CACAACAGGC TGGTATCCCG   20

(2) INFORMATION FOR SEQ ID NO: 19:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 20 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: cDNA

  (iii) HYPOTHETICAL: NO

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

CAATAACGAA CTGGGAAGCC   20

(2) INFORMATION FOR SEQ ID NO: 20:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 23 base pairs
       (B) TYPE: nucleic acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: cDNA

  (iii) HYPOTHETICAL: NO

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

ACTCACTATA GGGCTCGAGC GGC   23

(2) INFORMATION FOR SEQ ID NO: 21:

    (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 26 base pairs
       (B) TYPE: nucleic acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: cDNA

  (iii) HYPOTHETICAL: NO

   (ix) FEATURE:
       (A) NAME/KEY: modified_base
       (B) LOCATION: 4
       (D) OTHER INFORMATION: /mod_base= i

   (ix) FEATURE:
       (A) NAME/KEY: modified_base
       (B) LOCATION: 6
       (D) OTHER INFORMATION: /mod_base= i

   (ix) FEATURE:
       (A) NAME/KEY: modified_base
       (B) LOCATION: 9
       (D) OTHER INFORMATION: /mod_base= i

   (ix) FEATURE:
       (A) NAME/KEY: modified_base
       (B) LOCATION: 15
       (D) OTHER INFORMATION: /mod_base= i

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

GAYNTNATNT GGRTNCAYGA YTAYCA   26

(2) INFORMATION FOR SEQ ID NO: 22:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 20 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (iii) HYPOTHETICAL: NO

    (ix) FEATURE:
        (A) NAME/KEY: modified_base
        (B) LOCATION: 3
        (D) OTHER INFORMATION: /mod_base= i

    (ix) FEATURE:
        (A) NAME/KEY: modified_base
        (B) LOCATION: 6
        (D) OTHER INFORMATION: /mod_base= i

    (ix) FEATURE:
        (A) NAME/KEY: modified_base
        (B) LOCATION: 12
        (D) OTHER INFORMATION: /mod_base= i

    (ix) FEATURE:
        (A) NAME/KEY: modified_base
        (B) LOCATION: 18
        (D) OTHER INFORMATION: /mod_base= i

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

CCNACNGTRC ANGCRAANAC   20

(2) INFORMATION FOR SEQ ID NO: 23:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 28 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (iii) HYPOTHETICAL: NO

          (ix) FEATURE:
               (A) NAME/KEY: modified_base
               (B) LOCATION: 2
               (D) OTHER INFORMATION: /mod_base= i

5   (ix) FEATURE:
               (A) NAME/KEY: modified_base
               (B) LOCATION: 5
               (D) OTHER INFORMATION: /mod_base= i

10   (ix) FEATURE:
               (A) NAME/KEY: modified_base
               (B) LOCATION: 8
               (D) OTHER INFORMATION: /mod_base= i

15   (ix) FEATURE:
               (A) NAME/KEY: modified_base
               (B) LOCATION: 14
               (D) OTHER INFORMATION: /mod_base= i

          (ix) FEATURE:
20   (A) NAME/KEY: modified_base
               (B) LOCATION: 20
               (D) OTHER INFORMATION: /mod_base= i

25   (ix) FEATURE:
               (A) NAME/KEY: modified_base
               (B) LOCATION: 23
               (D) OTHER INFORMATION: /mod_base= i

30

          (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

35   TNGGNTKNTT YYTNCAYAYN CCNTTYCC   28

     (2) INFORMATION FOR SEQ ID NO: 24:

          (i) SEQUENCE CHARACTERISTICS:
40   (A) LENGTH: 20 base pairs
               (B) TYPE: nucleic acid
               (C) STRANDEDNESS: single
               (D) TOPOLOGY: linear

45   (ii) MOLECULE TYPE: cDNA

     (iii) HYPOTHETICAL: NO

50   (ix) FEATURE:
               (A) NAME/KEY: modified_base
               (B) LOCATION: 6
               (D) OTHER INFORMATION: /mod_base= i

```
            (ix)  FEATURE:
                  (A)  NAME/KEY: modified_base
                  (B)  LOCATION: 9
                  (D)  OTHER INFORMATION: /mod_base= i

      5
            (ix)  FEATURE:
                  (A)  NAME/KEY: modified_base
                  (B)  LOCATION: 18
                  (D)  OTHER INFORMATION: /mod_base= i

     10


            (xi)  SEQUENCE DESCRIPTION: SEQ ID NO: 24:


      TGRTCNARNA RYTCYTTNGC                                              20
     15
      (2) INFORMATION FOR SEQ ID NO: 25:


            (i)  SEQUENCE CHARACTERISTICS:
                 (A)  LENGTH: 20 base pairs
     20          (B)  TYPE: nucleic acid
                 (C)  STRANDEDNESS: single
                 (D)  TOPOLOGY: linear


           (ii)  MOLECULE TYPE: cDNA

     25
          (iii)  HYPOTHETICAL: NO


            (ix)  FEATURE:
                  (A)  NAME/KEY: modified_base
     30           (B)  LOCATION: 9
                  (D)  OTHER INFORMATION: /mod_base= i


            (ix)  FEATURE:
                  (A)  NAME/KEY: modified_base
     35           (B)  LOCATION: 12
                  (D)  OTHER INFORMATION: /mod_base= i


            (ix)  FEATURE:
                  (A)  NAME/KEY: modified_base
     40           (B)  LOCATION: 15
                  (D)  OTHER INFORMATION: /mod_base= i


            (ix)  FEATURE:
                  (A)  NAME/KEY: modified_base
     45           (B)  LOCATION: 18
                  (D)  OTHER INFORMATION: /mod_base= i



     50
            (xi)  SEQUENCE DESCRIPTION: SEQ ID NO: 25:

                                                                        20
      CCRTGYTCNG CNSWNARNCC
```

(2) INFORMATION FOR SEQ ID NO: 26:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 20 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: cDNA

  (iii) HYPOTHETICAL: NO

   (ix) FEATURE:
        (A) NAME/KEY: modified_base
        (B) LOCATION: 6
        (D) OTHER INFORMATION: /mod_base= i

   (ix) FEATURE:
        (A) NAME/KEY: modified_base
        (B) LOCATION: 17
        (D) OTHER INFORMATION: /mod_base= i

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

TCRTCNGTRA ARTCRTCNCC   20

(2) INFORMATION FOR SEQ ID NO: 27:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 22 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: cDNA

  (iii) HYPOTHETICAL: NO

   (ix) FEATURE:
        (A) NAME/KEY: modified_base
        (B) LOCATION: 3
        (D) OTHER INFORMATION: /mod_base= i

   (ix) FEATURE:
        (A) NAME/KEY: modified_base
        (B) LOCATION: 6
        (D) OTHER INFORMATION: /mod_base= i

   (ix) FEATURE:
        (A) NAME/KEY: modified_base
        (B) LOCATION: 15

(D) OTHER INFORMATION: /mod_base= i

     (ix) FEATURE:
          (A) NAME/KEY: modified_base
5   (B) LOCATION: 21
          (D) OTHER INFORMATION: /mod_base= i

10   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

     GYNACNARRT TCATNCCRTC NC   22

(D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
    (A) NAME/KEY: modified_base
5   (B) LOCATION: 21
    (D) OTHER INFORMATION: /mod_base= i

10   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

GYNACNARRT TCATNCCRTC NC

22